



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

Publication number:

0212914  
A2

B4 D16

12

# EUROPEAN PATENT APPLICATION

Application number: 86306073.7

Date of filing: 06.08.86

Int. Cl.: C 12 P 21/00, C 12 P 21/02,  
C 07 K 15/00, C 12 N 15/00,  
C 12 N 1/16

Priority: 06.08.85 US 763130

Date of publication of application: 04.03.87  
Bulletin 87/10

Designated Contracting States: AT BE CH DE FR GB IT LI  
LU NL SE

Applicant: IMMUNEX CORPORATION, 51 University  
Street Suite 600 600 Immunex Building, Seattle  
Washington 98101 (US)

Inventor: Deeley, Michael C., 1140 North 76th, Seattle  
Washington 98103 (US)  
Inventor: Gimpel, Steven D., 3602 North East 73rd, No. 5,  
Seattle Washington 98115 (US)  
Inventor: Price, Virginia L., 2617 Boyer Avenue East,  
Seattle Washington 98102 (US)

Representative: Sheard, Andrew Gregory et al, Kilburn &  
Strode 30, John Street, London WC1N 2DD (GB)

Amplifying the expression of recombinant DNA products.

Amplified expression of recombinant DNA products is achieved in hosts expressing protease that cleave at multi-basic amino acid residues. To this end, wild-type genes encoding the desired protein products are mutated by substituting codons or eliminating codons encoding multi-basic amino acid residues while maintaining the activity of the expressed protein product. Mutation of the desired gene can be conveniently carried out by site-specific in vitro mutagenesis.

B  
D  
Z

EP 0 212 914 A2

- 2637

87058395

AMPLIFYING EXPRESSION OF RECOMBINANT DNA  
PRODS. + BY USING HOSTS EXPRESSING  
PROTEASE THAT CLEAVES AT MULTI-BASIC  
AMINO ACID RESIDUES

## AMPLIFYING THE EXPRESSION OF RECOMBINANT DNA PRODUCTS

### Technical Field

The present invention relates to a method for amplifying the expression of recombinant DNA in hosts expressing protease that cleave at  
5 multibasic amino acid residues and to the use of this method in conjunction with colony stimulating factor (hereinafter "CSF") and, more particularly, human granulocyte-macrophage colony stimulating factor (hereinafter "GM-CSF").

### Background of the Invention

CSF refers to a family of lymphokines which induce progenitor  
10 cells found in the bone marrow to differentiate into specific types of mature blood cells. The particular type of mature blood cell that results from a progenitor cell depends upon the type of CSF present. For instance, erythropoietin is believed to cause progenitor cells to mature into erythrocytes, while thrombopoietin is thought to drive progenitor cells along the thrombocytic  
15 pathway. Similarly, granulocyte-macrophage colony formation is dependent on the presence of GM-CSF. The present invention concerns an analog of human GM-CSF.

CSF, including human GM-CSF, is produced only in minute quantities in vivo. CSF-like factors have been extracted from body organs, Sheridan  
20 and Stanley, J. Cell. Physiol. 78:451-459 (1971), and have been detected in serum and urine, Robinson et al., J. Cell. Physiol. 69:83-92 (1967); Stanley et al., J. Lab. Clin. Med. 79:657-668 (1972). Researchers have reported isolating low titer CSF-like factor from human peripheral blood cells which appear to be macrophages or monocytes, Moore and Williams, J. Cell. Physiol. 80:195-206  
25 (1972); Golde and Kline, J. Clin. Invest. 51:2981-2983 (1972); Moore et al., J. Natl. Cancer Inst. 50: 591-601 (1973).

Although the factors identified by the above researchers have been reported to be CSF, heretofore sufficient quantities of homogeneous human CSF, including GM-CSF, have not been available to thoroughly investigate its bio-  
30 chemistry and biology. The availability of adequate quantities of homogeneous  
2 6 3 8

87058395

human GM-CSF would be valuable in investigations and possible treatment of proliferative blood disorders, such as certain leukemias and anemias. Also, human GM-CSF in greater purity and larger quantities than heretofore available could prove useful in achieving successful bone marrow transplantation following cancer chemotherapy.

One potential method of providing larger quantities of homogeneous polypeptides including human GM-CSF than heretofore available is through recombinant DNA techniques. Recombinant DNA techniques have been developed for economically producing a desired protein once the gene coding for the protein has been isolated and identified. A discussion of such recombinant DNA techniques for protein production is set forth in the editorial and supporting papers in Vol. 196 of Science (April 1977).

#### Summary of the Invention

Genes encoding various protein products have been isolated and cloned for expression of functional protein product in yeast expression systems employing the promoter and leader sequence for the yeast pre-pro- $\alpha$  mating factor (" $\alpha$ -factor"). Although larger quantities of mature, homogeneous protein product were achieved through the yeast expression system than heretofore produced, applicants hypothesized that the level of protein product being recovered was possibly somehow being restricted by the existence of a potential cleavage site for the protease encoded by the KEX 2 gene of the yeast Saccharomyces cerevisiae ("S. cerevisiae"). This secretory pathway processing enzyme was found to cleave at "double basic amino acid residues," i.e., two adjacent basic amino acid residues located along the amino acid sequence of the protein product. In an attempt to increase the levels of mature protein product recovered in yeast systems, applicants sought to alter wild-type genes to eliminate "multibasic amino acid residues," i.e., two or more adjacent basic amino acid residues located along the amino acid sequence of the protein product, by substitution or deletion of codons encoding multibasic residues.

The present invention has been carried out with respect to GM-CSF. Different types of CSF, including GM-CSF have been discussed supra. Although a substantial portion of the remainder of the application discusses present invention with respect GM-CSF, it is to be understood that the present invention is not limited to GM-CSF, but rather may be employed in conjunction with virtually all protein products that naturally are composed of multibasic amino acids. In addition, the present invention is not limited to the use of yeast cells as hosts, but rather is applicable to any host that expresses a protease that

cleaves precursor protein products at double basic amino acid residues during the expression process.

An analog GM-CSF is produced by altering the wild-type gene for GM-CSF by replacing the applicable codons coding for basic amino acids to eliminate multibasic amino acids. One possible and preferred technique of making this substitution is by site-specific in vitro mutagenesis, for instance as discussed in Craik, Biotechniques, Jan. 1985, 12-19. In this procedure, the gene coding for human GM-CSF is ligated into an M13 single-stranded filamentous phage vector which is then employed to transform an appropriate host to produce replicate single-stranded DNA templates. The single-stranded DNA templates are annealed with portions of a complementary M13 strand to form a gapped heteroduplex. A synthesized mutagenesis oligonucleotide constructed with the altered/replaced codon, is annealed to the corresponding portion of the wild-type GM-CSF gene disposed in the single-stranded region of the gapped heteroduplex. The gaps between the ends of the mutation oligonucleotide and the complementary M13 strand are enzymatically repaired to form a double stranded structure which is then used to transform an appropriate host. Properly mutagenized genes are detected by use of a radiolabeled oligonucleotide probe having the same structure as the mutagenesis oligonucleotide.

Thereafter, the nucleotide sequence of candidates identified with the radiolabeled probe is determined to verify the desired gene construction had been achieved. Next, the altered gene is transferred from the M13 vector to a yeast expression vector used to transform S. cerevisiae for expression of mature, analog GM-CSF. Biological assays are conducted to confirm that the analog GM-CSF exhibited substantially the same activity as the natural GM-CSF product.

An analog GM-CSF is also produced by altering the wild-type gene encoding GM-CSF by removing the applicable codons encoding basic amino acids to eliminate the occurrence of multibasic amino acid residues. As an illustrative but non-limiting example, the applicable codons may be deleted from the wild-type gene by the same site-specific in vitro mutagenesis technique discussed above regarding the replacement of codons encoding basic amino acid residues. In this procedure, the composition of the synthesized mutagenesis oligonucleotide is the same as the corresponding portion of the wild-type gene, however with the applicable codon or codons encoding the basic amino acid(s) deleted. With this exception the procedures for preparing recombinant DNA encoding the analog GM-CSF, for expression of the analog product and for

87058395

biological assay to confirm the functionality of the analog GM-CSF is the same as discussed above.

#### Brief Description of the Drawings

The details of typical embodiments of the present invention will be described in connection with the accompanying drawings, in which:

FIGURE 1A illustrates the amino acid and nucleotide sequences of the wild-type human GM-CSF gene, including part of the 3' non-coding region of the gene;

FIGURE 1B illustrates the amino acid and nucleotide sequences of a mutant human GM-CSF gene wherein at least one codon coding for a basic amino acid residue has been replaced by a codon encoding for a nonbasic amino acid residue so that the analog GM-CSF encoded by the mutant gene is devoid of multibasic residues;

FIGURE 1C illustrates the amino acid and nucleotide sequences of a mutant human GM-CSF gene wherein at least one codon encoding a basic amino acid residue has been deleted so that the analog GM-CSF encoded by the mutant gene is devoid of multibasic amino acid residues;

FIGURE 2 illustrates the plasmid pYα fGM-2 used to direct expression of the wild type GM-CSF in yeast hosts;

FIGURE 3 illustrates the strategy employed for generating the mutated, codon substituted gene, M13HuGMLeu23, coding for analog human GM-CSF;

FIGURE 4 illustrates the pYα fHuGMLeu23 expression plasmid with the coding region for the mutated GM-CSF gene, M13HuGMLeu23, inserted therein for use in transforming host cells for amplified expression of functional analog human GM-CSF;

FIGURE 5 illustrates the strategy employed for generating the mutated, codon deleted gene M13HuGMΔ Arg23 encoding analog GM-CSF; and,

FIGURE 6 illustrates the pYα fHuGMΔ Arg23 expression plasmid with the coding region for the mutated GM-CSF gene, M13HuGMΔ Arg23, inserted therein for use in transforming host cells amplified expression of functional analog GM-CSF.

#### Description of Preferred Embodiments

##### Isolation of the Wild-Type Human GM-CSF Gene

The wild-type gene coding for human GM-CSF has been isolated and characterized. The nucleic acid sequence of the gene is shown in FIGURE 1. This wild-type gene, inserted into a cloning plasmid, designated as pHG23 and then transformed into *E. coli*, is on deposit with the American Type Culture

87058395

Collection ("ATCC"), 12301 Parklawn Drive, Rockville, Maryland 20852, USA, under Accession No. 39900. Also, the wild type gene as inserted into a yeast expression plasmid, designated as pY $\alpha$ fGM-2, as shown in FIGURE 2, is on deposit with the ATCC under Accession No. 53157.

5 In brief summary, the gene coding for wild-type human GM-CSF was isolated from a cDNA library with a nick-translated cDNA probe. The probe was isolated from a murine GM-CSF cDNA library by use of a synthetic oligonucleotide probe corresponding to a portion of the nucleotide sequence of murine GM-CSF. Total human RNA was extracted from the HUT-102 lymphoma  
10 T-cell line and from peripheral blood T-lymphoma cells, and then polyadenylated mRNA was isolated from the total RNA extract. A cDNA library was constructed by reverse transcription of the polyadenylated mRNA with the enzyme reverse transcriptase. The DNA was rendered double-stranded with DNA polymerase I and inserted into an appropriate cloning vector. Recombinant  
15 cloning vectors were used to transform an appropriate host.

Transformed hosts were identified and grouped into pools. Plasmid DNA prepared from these pools was hybridized with the murine cDNA probe that had been radiolabeled. The pool(s) of clones that gave a positive signal to the probe were identified and then the putative pools subdivided and a hybridization  
20 screen repeated. A single transformant corresponding to the wild-type human GM-CSF gene was eventually identified. Plasmid DNA was prepared from this transformant and characterized by DNA sequencing. The coding region of the wild-type human GM-CSF gene was employed to construct an expression plasmid designated as pY $\alpha$ fGM-2 and illustrated in FIGURE 2, for use in a yeast host  
25 system to express mature GM-CSF. The expression plasmid was constructed with the yeast pre-pro- $\alpha$  mating factor (" $\alpha$ -factor") as an efficient promoter together with leader sequences to direct the synthesis and secretion of GM-CSF in yeast. A synthesized oligonucleotide, defined in FIGURE 2, containing a 5' cohesive terminal and a second  $\alpha$ -factor processing site was coupled to the 5' end  
30 of the GM-CSF gene to facilitate plasmid construction and improve expression levels. Thereafter, biological assays were conducted which confirmed that the expressed protein product was GM-CSF. The assay ascertained the ability of the GM-CSF to direct the formation of mixed, granulocytic- and macrophage-type colonies from human bone marrow cells. The GM-CSF was found to direct  
35 synthesis of GM-CSF activity in the bone marrow colony assay at a level of approximately  $1.25 \times 10^6$  colony forming units ("CFU") per ml of culture supernatant.

87058395

Cloning of GM-CSF Gene Mutated by Codon Substitution

In accordance with one aspect of the present invention, a mutated human GM-CSF gene is prepared by replacing codons encoding basic amino acid residues with codons encoding nonbasic amino acid residues. The mutated gene  
5 is cloned for use in the expression of analog GM-CSF that is devoid of multibasic amino acid residues. In one specific form of the present invention, the codon encoding amino acid residue No. 23, arginine, in the wild-type gene has been substituted with a codon encoding a nonbasic amino acid residue thereby eliminating the multibasic sequence arginine-arginine at amino acid residues  
10 Nos. 23 and 24 of the GM-CSF, FIGURE 1A. The replacement residue may be composed of any nonbasic amino acid residues; however, the replacement residue chosen should not result in the creation of an enzyme cleavage site resulting in the undesirable cleavage of the GM-CSF expression product. Preferably, the replacement amino acid residue may include leucine or any other amino acid  
15 except lysine. Ideally, the replacement amino acid is composed of leucine.

It is to be understood that rather than replacing the arginine at amino acid residue No. 23 with a nonbasic residue, it is also within the scope of the present invention to instead replace the arginine at amino acid residue No. 24 with an appropriate nonbasic amino acid, for instance, with one of the  
20 amino acids set forth above. In addition, both arginine residues Nos. 23 and 24 could be replaced with nonbasic amino acid residues. An essential criteria regarding the particular amino acid residue(s) that is replaced is that the replacement results in the elimination of multibasic amino acids while substantially maintaining the biological activity of the GM-CSF.

Ideally the codon encoding of the nonbasic amino acid residue is chosen for maximum gene expression by host cells. It is known that in S. cerevisiae products encoded by genes composed of specific codon compositions are expressed more highly than products encoded by the same gene with an alternative codon composition for a particular amino acid residue. As a specific  
30 example, highly expressed genes in S. cerevisiae contain the TTG codon 92% of the time when encoding a leucine residue, and the other five leucine encoding codons only 8% of the time. Thus, in GM-CSF if the replacement residue is leucine, ideally the codon TTG will be employed.

The analog GM-CSF of the present invention preferably is produced by recombinant DNA methods employing a mutated GM-CSF gene coding  
35 for the analog protein product. In one preferred form of the present invention, the mutated gene is produced by substituting codon(s) encoding the desired nonbasic amino acid residue in place of codon(s) encoding the target basic amino

87058395

acid residue(s). Various site-specific mutagenesis procedures may be used for making this substitution including oligonucleotide-directed site-specific mutagenesis techniques, as discussed generally by Craik, supra. One method utilizes a synthetic oligonucleotide-defined sequence which is complementary to the region of the cloned DNA molecule except for the one to several desired nucleotide mismatches. The synthesized oligonucleotide is annealed with a single-stranded template clone (+) of the original (wild-type) DNA molecule carried in a phage vector. Even though the synthesized oligonucleotide does not perfectly correspond with the single-stranded template clone, it will anneal under proper (nonstringent) hybridization conditions, especially if the mismatches are located at or near the middle of the oligonucleotide rather than at one of the ends. The mismatched oligonucleotide serves as a primer for DNA polymerase to synthesize the remainder of the complementary (-) strand, resulting in a double-stranded molecule which is employed to transform an appropriate host for the repair of the mismatches and to produce both the wild-type and mutant genes.

As a somewhat modified and preferred technique, the single-stranded DNA template (+) can be annealed with portions of a complementary (-) phage strand together with the synthesized mutagenesis oligonucleotide, thereby leaving gaps between the ends of the oligonucleotide and the complementary (-) strand fragment. These gaps are enzymatically filled, and then the gap-filled duplexed DNA is transformed into an appropriate host for replication of the mutant gene.

Other site-specific mutagenesis techniques also may be employed in conjunction with the present invention to substitute for codons coding for multibasic amino acids in the GM-CSF gene. For instance, methods have been developed for generating single-stranded regions in double-stranded DNA molecules to allow annealing of a mutator oligonucleotide to the sequence of interest. One such technique involves making a single-stranded nick in the plasmid DNA with a restriction endonuclease in the presence of ethidium bromide and then extending the nick into a gap with Micrococcus luteus DNA polymerase. Shortle et al., Proc. Nat. Acad. Sci. (USA) 79:1588-1592 (1982). A mutated oligonucleotide can then be annealed to the single-stranded portion of the plasmid, and the gaps at the ends of the oligonucleotide enzymatically repaired.

As a further alternative "gapped duplexes" can be prepared from double-stranded DNA molecules by the controlled digestion of a nicked or linearized plasmid with exonuclease III. Wallace et al., Nucl. Acids Res.

87058395



9:3647-3658 (1981); and, Dalbadie-McFarland et al., Proc. Nat. Acad. Sci. (USA) 79:6409-6413 (1982).

#### Preparation of Single-Stranded DNA Template

• Single-stranded DNA templates corresponding to the wild-type GM-CSF gene are prepared by cloning the wild-type gene in phage vectors capable of producing single-stranded DNA molecule product when double-stranded replicative form DNA is used as a cloning vector. One such strain of phage is M13. See Hu and Messing, Gene, 17:271-277; and, Messing, Methods In Enzymology, 101:20-78 (1983). The replicative form DNA phage cloning vector preferably is constructed with duplexed oligonucleotides attached to the 5' terminal of the GM-CSF gene for use in linking the mutated GM-CSF gene to the  $\alpha$ -factor promoter and leader sequences contained in the expression plasmid employed to express the mutated GM-CSF gene, as discussed infra. An example of such duplexed oligonucleotides is shown in FIGURE 3. Ideally, the duplexed oligonucleotides together form a second  $\alpha$ -factor processing site at the 3' end of the oligonucleotide adjacent the 5' end of the GM-CSF gene to improve expression levels.

The phage vector, with the duplexed linking oligonucleotide and the wild-type GM-CSF gene inserted therein, is used to transfect an appropriate bacteria host, such as various strains of E. coli. Typical E. coli strains that may be used in conjunction with the present invention include strains JM101, JM103, JM105, and JM107 of E. coli K12 (Bethesda Research Laboratories, Bethesda, Maryland).

#### Preparation of Oligonucleotide

The oligonucleotide containing the desired codon substitution from the wild-type GM-CSF gene may be readily synthesized by well-known techniques, such as by phosphodiester or triester methods. The details of the triester synthesis technique are set forth, for example, in Sood et al., Nucl. Acid Res. 4:2557 (1977); and, Hirose et al., Tet. Lett. 28:2449 (1978).

Preferably, the substituted codon is located at approximately the center of the oligonucleotide, and the oligonucleotide is long enough to readily hybridize to the single-stranded DNA as prepared above, while being short enough to be relatively easily synthesized. As an illustrative but nonlimiting example, if, as discussed above, the arginine amino acid residue No. 23 of the wild type GM-CSF gene is substituted with leucine, then the oligonucleotide, designated MCD5-27, could be of the following composition: 5'-CATCCAG-GAGGCCTTGCGTCTCCTGAA-3'. In this oligonucleotide construction the codon corresponding to leucine, TTG, as underlined, is located near the center of the

oligonucleotide. As noted above, this particular composition of the codon encoding leucine was chosen to maximize analog GM-CSF expression. It is to be understood that a smaller number or larger number of flanking nucleotides may be employed, and that the substitute codon does not necessarily have to be located at this position of the oligonucleotide.

#### Cloning of Mutated Gene

Referring to FIGURE 3, for use in forming the heteroduplex DNA, double stranded wild-type M13 DNA ideally, but not necessarily, is prepared from the same strain used to form the single-stranded template. Preferably, the double stranded DNA overlaps substantially the entire template strand (+) except in the region of the substituted codon.

The wild-type M13 DNA portion and the oligonucleotide are annealed with the template strand (+) by well-known standard procedures to form the gapped duplex structure. The gaps between the ends of the oligonucleotide and the corresponding ends of the complementary (-) strand are filled in by standard techniques employing E. coli DNA polymerase ("Klenow" fragment) and T4 DNA ligase. Thereafter the covalently closed heteroduplex is employed to transform an appropriate host, such as a strain of E. coli. Upon transfection of the host and replication of the heteroduplex, mixed progeny containing either the wild-type or mutant copies of the GM-CSF gene are produced.

#### Screening of Cloned DNA Molecules

Plaques resulting from the transfection of the host are screened for the oligonucleotide-directed mutant DNA molecules with a radiolabeled oligonucleotide probe, ideally of the same composition as the mutation oligonucleotide. Although the oligonucleotide probe may be radiolabeled by many different techniques and with many different isotopes, of preference is the radiolabeling of the probe with T4 polynucleotide kinase and  $^{32}\text{P}$ -ATP. A standard protocol for the labeling procedure is set forth in Maniatis et al., Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring, New York (1982).

Putative plaques are picked and screened with the  $^{32}\text{P}$ -labeled oligonucleotide probe. The picked plaques are used to inoculate microtiter wells containing YT medium. After a suitable growth period, the candidate cultures are spotted onto nitrocellulose filters placed on YT plates. After further growth, the DNA is liberated and bound to the nitrocellulose filter. The bound DNA is then hybridized with the labeled oligonucleotide probe. The specific DNA fragments that hybridize to the probe are identified by autoradiography. By this procedure candidates containing the site-specific mutation are identified.

87058395

Single-stranded phage and double-stranded replicative form DNA containing the site-specific mutation, designated as M13HuGMLeu23, are prepared.

#### Characterization of Screened GM-CSF Mutation

The single-stranded phage DNA prepared above is sequenced using  
5 standard chain-termination methods. This technique of nucleotide sequencing was originated by Sanger et al., Proc. Natl. Acad. Sci. (USA) 70:5463 (1977). See U.S. Patent No. 4,322,499. Methods for chain-termination sequence determination are set forth in: the Amersham Handbook entitled, M13 Cloning and Sequencing, Blenheim Crescent, London (1983) (hereinafter "Amersham  
10 Handbook"); Messing, 2 Recombinant DNA Technical Bulletin, NIH Publication No. 79-99, 2 43-48 (1979); Norrander et al., Gene 26:101 (1983); Cerretti et al., Nucl. Acids Res. 11:2599 (1983); and, Biggin et al., Proc. Natl. Acad. Sci. (USA) 80:3963 (1983).

In the chain-termination sequencing method, single-stranded  
15 template molecules are primed with a short universal primer strand having a free 3' hydroxyl group and then using DNA polymerase (Klenow fragment) to copy the template strand in a chain extension reaction using all four deoxyribonucleotide triphosphates, i.e., dATP, dCTP, dGTP, and dTTP (collectively referred to as "dNTPs"), with one of the dNTPs being radiolabeled. In the synthesis reaction, a  
20 nucleotide specific chain terminator lacking a 3'-hydroxyl terminus, for instance, a 2', 3' dideoxynucleotide triphosphate ("ddNTP"), is used to produce a series of different length chain extensions. The terminator has a normal 5' terminus so that it can be incorporated into a growing DNA chain, but lacks a 3' hydroxyl terminus. Once the terminator has been integrated into a DNA chain, no further  
25 deoxynucleotide triphosphates can be added so that growth of the chain stops. Four separate synthesizing reactions are carried out, each having a ddNTP of one of the four nucleotide dNTPs, i.e., dATP, dCPT, dGTP and dTTP. One of the normal dNTPs is radiolabeled so that the synthesized strands, after having been sorted by size on a polyacrylamide gel, can be autoradiographed. The chain  
30 extensions from the four reactions are placed side by side in separate gel lanes so that the pattern of the fragments from the autoradiography corresponds to the nucleic acid sequence of the cloned DNA.

FIGURE 1B illustrates the nucleotide sequence of the mutated human GM-CSF gene contained in the M13HuGMLeu23 plasmid DNA. The  
35 corresponding amino acid composition of the coding region of the mutant gene is also illustrated in FIGURE 1B, beginning from the Ala residue, No. 1 (nucleotide No. 14) and extending to the Glu residue, No. 127 (nucleotide No. 394). As expected, the M13HuGMLeu23 mutant differed from the wild-type gene,

FIGURE 1A, only at the twenty-third codon in which the altered gene contained the sequence TTG (Leu) rather than CGG (Arg). In FIGURE 1B, the nucleotides 5' of the coding region of the mutant gene compose the second  $\alpha$ -factor processing site and a Hind III cohesive 5' terminal (nucleotide Nos. -6 to 13).

It is to be understood that rather than employing the chain-termination technique outlined above, other known methods may be utilized to sequence cloned human cDNA inserts without departing from the spirit or scope of the present invention. For instance, the chemical degradation method of Maxam and Gilbert as set forth in Proc. Nat'l Acad. Sci. (USA) 74:560 (1977) can be used.

#### Expression of Analog GM-CSF

The M13HuGMLeu23 cDNA fragment shown in FIGURE 1B, from the Hind III restriction site (nucleic acid No. -6) to Neo I restriction (nucleic acid No. 502) is inserted into an expression vector (see FIGURE 4) designed to direct synthesis and secretion of the mature form of analog GM-CSF from yeast host cells. The expression vector, for instance pY $\alpha$ fHuGMLeu23, preferably contains sequences derived from plasmid pBR 322 containing an origin of replication and the ampicillin resistance gene (Amp<sup>r</sup>) (thick line portion in FIGURE 4). Preferably, the expression vector also includes sequences from yeast, for instance the tryptophan-1 gene (Trp-1) as a selectable marker and the 2 u yeast origin of replication (thin line portion in FIGURE 3). Ideally, the expression vector further includes the yeast  $\alpha$ -factor (for instance, stippled box portion) as an efficient promoter together with leader sequences to direct the synthesis and secretion of GM-CSF in yeast hosts, followed by the second  $\alpha$ -factor processing site (open box portion) derived from the duplexed linking oligonucleotide and then the sequence for the coding region of GM-CSF (hatched box portion). The structure of the  $\alpha$ -factor gene is discussed in Kurjan and Herskowitz, Cell 30:933-943 (1982).

The pY $\alpha$ fHuGM Leu 23 expression plasmid is transformed into an appropriate strain of S. cerevisiae. Preferable strains include, but are not limited to, yeast strains 79, X2181-1B, DBY746, YNN282, 20B-12. These strains are all  $\alpha$ , Trp 1 for compatibility with the  $\alpha$ -factor promoter and for selection of Trp<sup>+</sup> transformants. These strains are all widely available, for instance strain 79 is available from the Yeast Genetic Stock Center, Department of Biophysics and Medical Physics, University of California, Berkeley, California 94702.

87058395

2648

Transformation of the yeast host with the recombinant expression plasmid containing the mutated GM-CSF gene is conducted according to well known procedures wherein spheroplasts are formed and then washed prior to plasmid uptake. Standard protocols for this procedure have been established.

- 5 See Beggs, Nature (London) 275:104 (1978); Hinnen et al., Proc. Natl. Acad. Sci. (USA) 75:1929 (1978).

The yeast culture supernatants are assayed for biological activity through their ability to direct the formation of mixed, granulocytic and macrophage-type colonies from human bone marrow cells. As a control,  
10 plasmid pY $\alpha$ f, of the same construction as pY $\alpha$ fHuGMLeu23 but lacking the GM-CSF sequences, was also transformed into a yeast host and the culture supernatant tested for biological activity. The pY $\alpha$ fHuGMLeu23 supernatant was found to direct synthesis of high levels of GM-CSF activity in the bone marrow colony assay ( $7.2 \times 10^6$  CFU-C/ml): whereas, no activity was detected  
15 from the supernatant derived from the pY $\alpha$ f control plasmid.

Cloning, Screening and Characterization of GM-CSF Gene Mutated by Codon Deletion and Expression of Analog GM-CSF With Mutated Gene

In accordance with another aspect of the present invention,, a mutated human GM-CSF gene is prepared by deleting codons encoding basic  
20 amino acid residues, and the mutated gene is cloned for expression of analog GM-CSF that is devoid of multibasic amino acid residues. In one specific form of this aspect of the present invention, the codon encoding amino acid residue No. 23, arginine, in the wild-type gene is deleted thereby eliminating the multibasic sequence arginine-arginine at amino acid residue Nos. 23 and 24 of  
25 the GM-CSF, FIGURE 1A. It is to be understood that rather than deleting the arginine at amino acid residue No. 23, it is also within the scope of the present invention to delete the arginine at amino acid residue No. 24, or to delete both arginine residues at residue Nos. 23 and 24. An essential criterion regarding the particular amino acid residue(s) deleted is that the deletion results in the  
30 elimination of multibasic amino acids while maintaining the biological activity of the GM-CSF.

As in the above-discussed GM-CSF gene mutated by codon substitution, the analog GM-CSF employing the GM-CSF gene mutated by codon deletion preferably is produced by recombinant DNA methods. In one preferred  
35 form of the present invention, the mutated gene having codon(s) deleted is produced by the same site-specific mutagenesis technique discussed above relative to the GM-CSF gene mutated by codon substitution, with the exception that rather than synthesizing the oligonucleotide with a substituted, nonbasic amino acid codon, the oligonucleotide is prepared by deletion of one or more of  
2649

87058395

the target codons encoding basic amino acid residues deleted. With this one exception, the same procedure discussed above for preparing the mutant gene by codon substitution is used to prepare the mutant gene by codon deletion. In the procedure the single-stranded DNA template (+) corresponding to the wild-type GM-CSF gene is prepared by cloning the wild-type gene in phage vectors capable of producing single-stranded DNA product when double-stranded replicative form DNA is used as a cloning vector. The single-stranded DNA template is annealed with portions of a complimentary (-) phage stranded together with the synthesized, codon deleted mutagenesis oligonucleotide thereby leaving gaps between the ends of the oligonucleotide and the complimentary (-) strand fragment. These gaps are enzymatically filled and then the gap-filled duplexed DNA is transformed into an appropriate host for replication of the codon deleted mutant gene.

As discussed above, the oligonucleotide containing the desired codon deletion from the wild-type GM-CSF gene is synthesized by well-known techniques, such as by phosphodiester or triester methods. Preferably the codon deletion is located at approximately the center of the oligonucleotide, and the oligonucleotide is long enough to readily hybridize to the single-stranded DNA template while being short enough to be relatively easily synthesized. As an illustrative but nonlimiting example, if, as discussed above, the arginine amino acid residue No. 23 of the wild-type GM-CSF gene is deleted, then the oligonucleotide, designated as MCD5-24, could be of the following composition: 5'-CATCCAGGAGGCCCGTCTCCTGAA-3'. It is to be understood that a smaller or larger number of flanking nucleotides may be employed on either side of the location of the deleted codon.

The GM-CSF gene mutated by codon deletion is cloned and screened in the same manner discussed above relative to the gene mutated by codon substitution. Also, plasmid DNA designated as M13HuGM $\Delta$ Arg23, containing the deleted arginine residue at position No. 23, is sequenced using standard chain-termination methods as discussed above beginning at page 9. FIGURE 1C illustrates the nucleotide sequence of the mutated GM-CSF gene contained in the M13HuGM $\Delta$ Arg23 plasmid DNA. The corresponding amino acid composition of the coding region of the mutated gene is also illustrated in FIGURE 1C, beginning with Ala residue (No. 1) (nucleotide No. 14) and extending to the Glu residue, N . 126 (nucleotide N . 391). As shown in FIGURE 1C, the M13HuGM $\Delta$ Arg23 differs from the wild-type gene, FIGURE 1A, only at the 23rd codon which was missing in the altered gene. The nucleotides 5' of the coding region of the mutant gene are composed of the  $\alpha$ -factor processing site

and a Hind III cohesive 5' terminal, as illustrated in FIGURE 1C (nucleotide Nos. -6 to 13).

Analog GM-CSF is expressed with the mutant M13HuGMΔArg23 gene using the same procedure employed to express analog GM-CSF using the M13HuGMLeu23 mutant gene discussed above beginning at page 11. Also the expressed protein product is tested for biological activity using the same bone marrow assay discussed above.

The processes and products of the present invention are further illustrated by the following examples.

10

### EXAMPLE I

#### Preparation of Single-Stranded DNA Template

As shown in FIGURE 3, a 487 base pair DNA fragment containing the coding region and a portion of the 3' flanking region of the human GM-CSF gene (extending from nucleotide No. 16 to nucleotide No. 502 in FIGURE 1A) was isolated from the pHG23 plasmid by digestion with the restriction enzymes Sfa NI and Nco I. T4 DNA polymerase was employed to blunt end the Nco I site of the gene fragment. Digestion of the pHG23 plasmid with the Sfa NI results in elimination of the first two nucleotides of the coding region of the GM-CSF gene. A duplexed linking oligonucleotide of the composition set forth below in Table 1 was synthesized to add back the two nucleotides of the initial Ala amino acid and also provides a second α-factor processing site for use in subsequent high level expression of the mutated GM-CSF gene in yeast hosts, as discussed more fully in Example 6. As shown in Table 1, the duplexed oligonucleotide is constructed with a cohesive Hind III 5' terminal.

25

TABLE 1

5'-	AGCT TCT	TTG	GAT	AAA	AGA	GC	-3'
3'-		AGA	AAC	CTA	TTT	TCT	CGT GGG -5'
		Ser	Leu	Asp	Lys	Arg	Ala Pro

- 2651

Hind III

↑  
factor  
processing

87058395

The isolated GM-CSF gene fragment, together with the individual oligonucleotides composing the duplex shown in Table 1 are ligated into the strain mp10 of the M13 phage vector (Amersham, Arlington Heights, Illinois), which was previously digested with the Hind III and Sma I restriction enzymes.

5 Ligation was accomplished in a reaction mixture composed of 20 nanograms (ng) of linearized mp10M13, 50 ng of the mutated GM-CSF gene fragment, 5 ng of synthetic oligonucleotides, one unit of T4 DNA ligase and sufficient T4 ligase buffer (0.4 M Tris [ph 7.4], 0.1 M MgCl<sub>2</sub>, 0.1 M dithiothreitol, 10 mM spermidine, 10 mM ATP and 1 mg/microliter ("ul") BSA) to form a 20 ul reaction

10 volume. Reaction was carried out by incubation at 25°C for 15 hours.

The M13mp10 vector with the DNA fragment inserted therein, designated as M13HuGM, was used to transfect by standard protocol *E. coli*. JM103 of the strain K12 (Bethesda Research Laboratories, Bethesda, Maryland) to produce a strain of *E. coli* actively excreting M13HuGM phage containing

15 single strand DNA. The phage were harvested from the culture supernatant after four hours of growth at 37°C by precipitation with polyethylene glycol. Single stranded DNA was isolated from the phage by extraction with phenol:chloroform according to standard protocol as detailed in the Amersham Handbook.

20

### EXAMPLE 2

#### Oligonucleotide Synthesis and Radiolabeling

The oligonucleotide employed for site-directed mutagenesis of the GM-CSF gene by codon substitution was chemically synthesized by standard triester method, as detailed by Sood et al., *supra* and Hirose et al. *supra*. The

25 oligonucleotide, designated as MCD5-27, was composed of the following sequence: 5'-CATCCAGGAGGCCTTGCGTCTCCTGAA-3'. The oligonucleotide was deblocked and purified by Sephadex G-50 chromatography (Pharmacia Fine Chemicals) followed by preparative gel electrophoresis.

The oligonucleotide was terminally radiolabeled with <sup>32</sup>P for use as

30 a screening probe. To facilitate radiolabeling, the 5' ends of the oligonucleotides were synthesized with OH termini, thereby eliminating the phosphatase treatment, which typically must be employed when labeling DNA fragments. The labeling protocol included adding 100 ng in 1 ul volume of the synthetic oligonucleotide to 16 ul of <sup>32</sup>P-ATP (7000 Ci/mM), 1 ul (10 U) of T4

35 polynucleotide kinase and 2 ul of 10 x kinase buffer I (0.5 M Tris-Cl [pH 7.0] 0.1 mM MgCl<sub>2</sub>, 50 mM dithiothreitol, 1 mM spermidine and 1 mM EDTA). The reaction is carried out at 37°C for 30 minutes, and then thereafter the <sup>32</sup>P



labeled oligonucleotides and unincorporated  $^{32}\text{P}$ -ATP were separated by Sephadex G-50 chromatography (Pharmacia Fine Chemicals).

### EXAMPLE 3

#### Site-Directed Mutagenesis of GM-CSF by Codon Substitution

5 As illustrated in FIGURE 3, for use in forming the gapped heteroduplex structure, strain mp18 of the M13 phage vector was digested with the restriction enzymes EcoRI and Hind III by standard techniques. The resulting major fragment and mutagenesis oligonucleotide MCD5-27 were annealed to the single-stranded template M13HuGM containing the wild type GM-CSF gene by  
10 the following procedure. One microgram ("ug") of the digested M13mp18 in double-stranded form was mixed with 0.5 ug of the single-stranded template DNA M13HuGM in 30 ul of 100 mM NaCl, 40 mM Tris-HCl (pH 7.5), 20 mM  $\text{MgCl}_2$ , 2.0 mM  $\beta$ -mercaptoethanol. The single-stranded template-double-stranded form M13mp18 fragment mixture was denatured by heating to 100°C for  
15 3 minutes and allowed to cool over 20 minutes to 65°C. Oligonucleotide MCD5-27 containing a 5'-phosphate (50.0 pmoles) was added and the mixture cooled slowly to 30°C and then placed on ice for 15 minutes. Thereafter the following were added to the mixture: 70 ul of 22 mM Tris-HCl (pH 7.5), 11 mM  $\text{MgCl}_2$ , 1.0 mM  $\beta$ -mercaptoethanol, 0.83 mM dATP, 0.83 mM dCTP, 0.83 mM  
20 dGTP, 0.83 mM dTTP, 0.4 mM rATP, 0.5 units of *E. coli* DNA polymerase (Klenow fragment) (Boehringer Mannheim Biochemicals), and 0.5 units T4 DNA ligase (Bethesda Research Laboratories). After an additional 30 minutes at 0°C, this primary extension mixture was incubated at 14.5°C for 20 hours.

### EXAMPLE 4

#### 25 Screening for Mutated Gene

The gap-filled duplex structure from EXAMPLE 3 was employed to transfect competent JM105 *E. coli* cells (Bethesda Research Laboratories, Bethesda, Maryland) by standard techniques, such as set forth in the Amersham Handbook, *supra*. The transfected JM105 cells were plated immediately after  
30 heat shock onto fresh YT plates in top agar.

Ninety-four of the resulting plaques were picked and screened with the radiolabeled MCD5-27 oligonucleotide probe, as prepared in EXAMPLE 2. The recombinant (white) plaques were picked with a sterile loop and used to inoculate microtiter dish wells containing 100 ul of YT medium. After about  
35 5-7 hours growth at 37°C a 96 well replicator was used to spot the candidate cultures onto nitrocellulose filters placed on YT plates, in duplicate. After overnight growth at 37°C, the filters were removed from the petri dishes. The DNA was liberated using alkali and neutralizing solutions by the general method

as described by Maniatus et al. supra. After the transfer process, the filter was air dried and baked for 2 hours at approximately 80°C to bind the single-stranded DNA to the nitrocellulose.

The bound DNA was next hybridized with the labeled oligo-  
5 nucleotide probe. Briefly, the baked nitrocellulose was incubated at 68°C for 2-4 hours in prehybridization buffer composed of: 6X standard saline-citrate ("SSC") (1X SSC is 0.15 M NaCl, 0.015 M NaCitrate, pH 7.0); and, 5X Denhardt's solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.2% bovine serum albumin). The filter was then incubated for 16 hours at 55°C with the <sup>32</sup>P-labeled  
10 oligonucleotide probe (10<sup>6</sup> cpm/ml, from EXAMPLE 2) in hybridizing buffer as above. After hybridization, the filter was washed extensively under high stringency conditions first with 6X SSC at room temperature and then for 1 hour at 68°C in 0.6X SSC. After air drying, the filter was subjected to auto-radiography at -70°C. This procedure resulted in clear identification of  
15 candidates containing the mutant GM-CSF gene, designated as M13HuGMLeu23.

#### EXAMPLE 5

##### Characterization of the Screened Mutagenized Gene

DNA templates were prepared from the candidates identified in  
EXAMPLE 4 and sequenced by standard chain-termination method as described  
20 in the Amersham Handbook, supra. The synthetic universal primer: 5'-CCCAGTCACGACGTT-3' (P-L Biochemicals, Milwaukee, WI), was annealed to the single-strand DNA templates and used to prime DNA synthesis as described above at page 10. Thereafter, the extension fragments were size-separated by gel electrophoresis and autoradiographed from which the nucleotide  
25 sequences of the fragments were deduced.

Deoxyadenosine 5' ( $\alpha$  [<sup>35</sup>S] thio) triphosphate (hereinafter "dATP [ $\alpha$  -<sup>35</sup>S]") was used as the radioactive label in the dideoxy sequencing reactions. Also, rather than using the gel set forth at page 36 of the Amersham Handbook, a 6% polyacrylamide gel was employed (6% polyacrylamide gel, 0.4 mm thick,  
30 containing 7 M, urea 100 mM Tris borate [pH 8.1], and 2 mM EDTA).

As noted above, the nucleotide sequence of the mutant GM-CSF gene, M13HuGMLeu23, is illustrated in FIGURE 1B, with the mature protein beginning at the asterisk (\*). The corresponding amino acid composition of the mature protein is set forth below the corresponding codons, beginning from the  
35 Ala residue, No. 1 (nucleotide No. 14) and extending to the Glu residue, No. 127 (nucleotide No. 394). As expected, this mutant gene differed from the wild type gene only at the 23rd codon in which the mutant gene contains the sequence TTG coding for leucine rather than the sequence CGT, coding for arginine. The

87058395

nucleotides 5' from the mature gene constitute the second  $\alpha$ -factor processing site and a Hind III cohesive 5' terminal, as illustrated in FIGURE 3.

#### EXAMPLE 6

##### Expression of Analog GM-CSF

5           The coding region, together with a portion of the 3' flanking region of the mutant GM-CSF gene as set forth in FIGURE 1B (nucleotides 14 through 502), and also together with the second  $\alpha$ -factor processing site and 5' Hind III cohesive terminal as coupled to the 5' end of the analog GM-CSF gene, was removed from the mutagenesis vector and employed to form a recombinant  
10 expression plasmid, designated as pY $\alpha$  fHuGMLeu23 to direct analog GM-CSF expression in yeast host cells. As shown in FIGURE 4, the pY $\alpha$  fHuGMLeu23 expression plasmid includes an origin of replication and an ampicillin resistant gene from plasmid pBR322 (thick line portion). The expression plasmid also includes the yeast 2 $\mu$  circle origin of replication and the Trp I gene for selection  
15 of transformed yeast hosts (TRP-[Trp-auxotrophs], (thin line portion in FIGURE 4). The expression plasmid further includes the  $\alpha$ -factor promoter and leader sequences used to direct transcription and secretion of the analog GM-CSF (stippled box portion). The analog GM-CSF sequences are shown in hatched box portion in FIGURE 4, whereas the Cathepsin B-like maturation site functioning as a second  $\alpha$ -factor processing site is shown as an open box portion  
20 in FIGURE 4, with the sequence thereof also set forth in FIGURE 4.

          The 5' leader sequence, the coding region and a portion of the 3' flanking region of the mutant GM-CSF gene, from nucleotide Nos. -6 to 502, was removed from the mutagenesis vector M13HuGMLeu23 and inserted into the  
25 expression vector pY $\alpha$  fHuGMLeu23 by digestion with the restriction enzyme Nco I followed by treatment with T4 DNA polymerase to fill in the recessed 3' end at the Nco I site with deoxynucleotides. Next the cleaved vector was treated with Hind III and the resulting 508bp M13HuGMLeu23 DNA fragment with the 5' leader sequence attached thereto was isolated by gel electrophoresis.  
30 This DNA fragment was ligated into the pY $\alpha$  f vector which previously had been prepared by removal of the Hind III-Stu I/Nco I section from the pY $\alpha$  fGM-2 expression plasmid (FIGURE 2) (ATCC No. 53157) by standard techniques.

          The pY $\alpha$  fHuGMLeu23 expression plasmid was transformed into yeast strain 79 ( $\alpha$ , Trp 1-1, Leu 2-1) of *S. cerevisiae* for selection of Trp<sup>+</sup>  
35 transformants by standard techniques. Prior to transformation, the strain 79 was grown in culture in YEPD medium (1% [wt/vol] yeast extract, 2% [wt/vol] peptone, 2% [wt/vol] glucose), to a density of  $2 \times 10^7$  cells/ml. Cells were

harvested by centrifugation at 1000 x g for 5 minutes at 22°C, and then the resulting pellet was washed with sterile, distilled water.

The yeast cells were then concentrated by resuspending in 1/10 vol. of SED (1 M sorbitol, 25 mM EDTA [pH 8.0], and 50 mM dithiothreitol) and incubating for 10 minutes at 30°C. The cell-buffer mixture was then centrifuged for 5 minutes at 300 x g. The pellet was washed once with 1/10 vol. of 1 M sorbitol and the cells resuspended in 20 milliliters of SCE (1 M sorbitol, 0.1 M sodium citrate [pH 5.8], 0.01 M EDTA). Glusulase, to break down the cell walls, in an amount of  $10^{-3}$  vol., was added to the solution and then the solution incubated at 30°C for 30 minutes with occasional gentle shaking. The presence of spheroplasts was assayed by diluting 10 microliters of the yeast cells into a drop of 5% SDS (wt./vol.) on a microscope slide to observe for "ghosts" at 400 x phase contrast. The cell mixture was then centrifuged at 300 x g for 3 minutes. The resulting pellet was twice washed with 1/10 vol. of 1 M sorbitol. The pellet was then once washed in CaS (1 M sorbitol, 10 mM  $\text{CaCl}_2$ ).

The yeast spheroplasts were then transformed with the previously prepared expression vector in a procedure adapted from Beggs, supra. The pelleted spheroplasts were suspended in 1/200 vol. of CaS and then divided into 100 microliter aliquotes in 1.5 ml Eppendorf tubes. Then, from 1 to 10 ul of the plasmid DNA were added to each aliquot (0.5 to 5 ug). The mixture was incubated at room temperature for 10 minutes and then 1 ml of PEG (20% PEG 4000, 10 mM  $\text{CaCl}_2$ , 10 mM Tris-HCl [pH 7.4]) was added to each aliquot to promote DNA uptake. After 10 minutes at room temperature, the mixture was centrifuged for 5 minutes at 350 x g. The resulting pellet was resuspended in 150 ul of SOS (10 ml of 2 M sorbitol, 6.7 ml of YEPD medium, 0.13 ml of 1 M  $\text{CaCl}_2$ , 27 ul of 1% tryptophan and 3.7 ml of water). This mixture was incubated for 20 minutes at 30°C. The cells were then plated.

Prior to plating the protoplast/DNA mixture, selective plates were preincubated at 37°C. Three ml of melted top agar (45°C), composed of 18.2 ml of sorbitol, 2 gm agar, 0.6 gm Difco yeast nitrogen base (without amino acids), 2 gm glucose, 0.1 ml of 1% adenine, 0.4 ml of 1% uracil and amino acids as required, was then added to each aliquot of transformed cells and the tube contents poured on the selective plates. The plates were incubated from 2 to 4 days at 30°C. Colonies which developed in the Trp minus medium contained plasmids that have the Trp 1 gene, i.e., those that are transformed.

Prior to biological assay, the transformants were grown in 20-50 ml of YEPD at 30°C to stationary phase. At the time of harvest, the protease inhibitors phenyl methyl sulfonyl flouride (PMSF) and Pepstatin A were

added to a final concentration of 1 mM and 10 uM, respectively. The cells were then removed by centrifugation at 400 x g and the medium was filtered through a 0.45 micron cellulose acetate filter.

#### EXAMPLE 7

##### 5 Colony Assay

The presence of analog GM-CSF harvested from the yeast cultures in Example 6 was confirmed by assaying the ability of the supernatant to stimulate growth of human bone marrow colonies in agar. For use in the assay, human bone marrow from the iliac crest of healthy donors was collected in a  
10 heparinized syringe. The marrow was diluted 1:3 with phosphate buffered saline (PBS) at room temperature and layered onto a solution of 54% percoll (Pharmacia Fine Chemicals). After centrifugation at 500 x g at room temperature for 20 minutes, the interface was collected and washed with  
20 volumes of PBS. The suspension was then centrifuged at 250 x g for 10  
15 minutes at room temperature. The cells were then resuspended in 10 ml of  $\alpha$ -Minimal Essential Medium with nucleotides ( $\alpha$ -Mem, Gibco) for cell counting and viability determination. FCS was then added and the cell suspension stored on ice until the assay was carried out.

In the assay, bone marrow cells as prepared above were added at a  
20 final concentration of  $1 \times 10^5$ /ml to an incubation medium consisting of: (a) seven parts of a solution containing 28.1% FCS,  $0.7 \times 10^{-4}$  M 2-mercaptoethanol, 0.12 mg/ml asparagine, 0.7 mg/ml glutamine, 150 units of penicillin G, 150 units of streptomycin,  $1.1 \times \alpha$ -MEM with nucleotides, and  $2.2 \times$  vitamins (Gibco); and, (b) three parts of 1.4% bacto-agar solution (Difco). The cultures  
25 were incubated in a humidified atmosphere at 37°C in the presence of 5% CO<sub>2</sub>. After seven to fourteen days of culture, the number and types of colonies, whether granulocyte, macrophage or mixed granulocyte-macrophage, were determined. Applicants found that the analog GM-CSF gene from the pY $\alpha$  fHuGMLeu23 clones directed synthesis of GM-CSF activity at the high level  
30 of  $7.2 \times 10^6$  colony forming units ("CFU") per milliliter. This activity level was determined by multiplying by 50 the reciprocal of the dilution giving 50% of the maximum colony number. Applicants have found that the average number of colonies from  $1 \times 10^5$  bone marrow cells was  $73 \pm 16$ . The colonies formed at 14 days by the recombinant GM-CSF were well defined and consisted of three types:  
35 approximately 1/3 mixed granulocyte-macrophage colonies; approximately 1/3 tight granulocyte colonies, and approximately 1/3 dispersed macrophage colonies.

87058395

As a control for the expression system of the present invention, a plasmid identical to pY $\alpha$  fHuGMLeu23, but lacking the GM-CSF sequences, was also transformed into yeast strain 79. The culture supernatant from the yeast produced no GM-CSF activity in the bone marrow colony assay.

5

#### EXAMPLE 8

##### GM-CSF Gene Mutated by Codon Deletion

The oligonucleotides employed for site-directed mutagenesis of the GM-CSF gene by codon substitution is chemically synthesized by standard triester method, as detailed by Sood et al. supra and Hirose et al. supra. The oligonucleotide, designated as MCD5-24, is composed of the following sequence:  
5'-CATCCAGGAGGCCCGTCTCCTGAA-3'. The oligonucleotide is deblocked and purified by Sephadex G50 chromatography (Pharmacia Fine Chemicals) followed by preparative gel electrophoresis. Thereafter, the oligonucleotide is terminally radiolabelled with  $^{32}\text{P}$  for use as a screening probe using the procedure discussed  
above in Example 2.

As illustrated in FIGURE 5, the MCD5-24 oligonucleotide is employed together with the single-stranded template containing the wild-type GM-CSF gene prepared in Example 1 and with the M13HuGM phage vector from Example 3 above to produce a gapped heteroduplex structure similar to that illustrated in FIGURE 3 by use of the procedures set forth in Example 3. Thereafter, the gap-filled duplex structure is employed to transfect competent JM105 E. coli cells (Bethesda Research Laboratories, Bethesda, MD) by standard techniques, as set forth in the Amersham Handbook, supra. Screening for the mutated gene in the transfected JM105 cells is carried out using the procedure set forth in Example 4, and the nucleic acid sequence of the screened mutated gene, designated as M13HuGM  $\Delta$  Arg23 is ascertained using the chain-termination method set forth in Example 5.

Analog GM-CSF is expressed using the procedure set forth in Example 6 wherein the 5' leader sequence, the coding region and a portion of the 3' flanking region of the mutant GM-CSF gene (from nucleotide members -6 to 502, is removed from the mutagenesis vector M13HuGM  $\Delta$  Arg23 by digestion with a restriction enzyme Nco I followed by treatment with T4 DNA polymerase and then cleavage with Hind III. The resulting 505 bp M13HuGM  $\Delta$  Arg23 DNA fragment with the 5' leader sequence attached thereto is isolated by gel electrophoresis and then ligated into the pY $\alpha$  f vector prepared by removal of the Hind 3- Stu I/Nco I section from the pY $\alpha$  fGM-2 expression plasmid (Figure 5) (ATCC No. 53157) by standard techniques. The resulting pY $\alpha$  fHuGM  $\Delta$  Arg23 expression plasmid (FIGURE 6) is transformed into yeast strain 79 as detailed in

Example 6 and then the expressed recombinant GM-CSF product tested for biological activity using the bone marrow colony assay set forth in Example 7.

As will be apparent to those skilled in the art in which the invention is addressed, the present invention may be embodied in forms other than those specifically disclosed above without departing from the spirit or essential characteristics of the invention. The particular embodiments of the present invention, described above, are therefore to be considered in all respects as illustrative and not restrictive. The scope of the present invention is as set forth in the appended claims rather than being limited to the examples contained in the foregoing description.

2659

87058395

CLAIMS

1. A biologically active analog polypeptide having an amino acid sequence devoid of multibasic amino acid residues.
2. A biologically functional analog polypeptide which is an analog of a biologically functional polypeptide having adjacent basic amino acid residues, the analog polypeptide being characterised by being free of adjacent basic amino acid residues.
3. An analog polypeptide as claimed in Claim 1 or 2, wherein multibasic amino acid residues in the corresponding naturally occurring form of the polypeptide have been eliminated by:
- (a) replacement of one or more basic amino acid residues with one or more non-basic amino acid residues; or,
  - (b) deletion of one or more basic amino acid residues.
4. An analog polypeptide as claimed in any one of Claims 1, 2 or 3, comprising granulocyte-macrophage colony stimulating factor.
5. An analog polypeptide as claimed in Claim 4, wherein either or both amino acid residues at position Nos. 23 or 24 shown in FIGURE 1A have been replaced with any amino acid residue except lysine or have been eliminated.
6. Recombinant or synthetic DNA coding for a

2660

87058395



polypeptide which normally has adjacent basic amino acid residues, the DNA being characterised by the codons coding for the normally adjacent basic amino acid residues being modified so that no two adjacent amino acid residues of the polypeptide coded for by the recombinant or synthetic DNA are both basic.

7. Substantially pure DNA encoding an analog polypeptide as claimed in any one of Claims 1 to 5.

8. A recombinant DNA vector for expression or cloning of DNA as claimed in Claim 7.

9. A host transformed by a vector as claimed in Claim 8.

10. A host as claimed in Claim 9, characterised by expressing a protease that cleaves at multibasic amino acid residues.

11. A method of amplifying the expression of recombinant DNA in a host that expresses protease that cleaves at multibasic amino acid residues, comprising the step of transfecting the host with recombinant DNA encoding a biologically active polypeptide devoid of multibasic amino acid residues.

12. A method as claimed in Claim 11, wherein the host is characterised by expressing a protease that cleaves at double basic amino acid residues.

13. A method as claimed in Claim 11 or 12, wherein the host is characterised by being of a strain of yeast.

14. A method as claimed in any one of Claims 11, 12 and 13, wherein the recombinant DNA encodes analog granulocyte-macrophage colony stimulating factor, with at least one basic amino acid residue in the naturally occurring granulocyte-macrophage colony stimulating factor being:

5 replaced with a non-basic amino acid residue to eliminate multibasic amino acid residues; and/or,  
deleted to eliminate multibasic amino acid  
10 residues.

15. A method as claimed in any one of Claims 9, 10 and 11, further comprising the step of preparing recombinant DNA encoding a polypeptide devoid of multibasic amino acid residues by:

15 replacing at least one codon encoding a basic amino acid residue in the DNA encoding the naturally occurring form of the polypeptide with a codon encoding a non-basic amino acid residue; or,

20 deleting from the DNA encoding the naturally occurring form of the polypeptide at least one codon encoding a basic amino acid residue.

16. A method as claimed in any one of Claims 11, 12, 13 and 15, wherein the recombinant DNA encodes analog granulocyte-macrophage colony stimulating factor.

17. A method as claimed in Claim 16, wherein the arginine amino acid residue at either or both position Nos. 23 or 24 shown in FIGURE 1A has been either replaced with one or more non-basic amino acid residues or deleted.

87058395

**Fig. 1A.** <sup>10</sup> \* <sup>30</sup> <sup>50</sup> <sup>70</sup>  
 CTGCAGCATCTCTGACCCGCCGCTGCCAGCCCCAGCACAACAGCCCTGGGAGCATGTGAAATGCCATC  
 CysSerIleSerAlaProAlaArgSerProSerProSerThrGlnProTrpGluHisValAsnAlaIle

<sup>90</sup> <sup>110</sup> <sup>130</sup> <sup>150</sup>  
 CAGGAGCCCGCTCTCTGAACCTGAGTAGAGACACTGCTGCTGAGATGAATGAACAGTAGAAGTCATCTCAGAAATG  
 GlnGluAlaArgArgLeuLeuAsnLeuSerArgAspThrAlaAlaGluMetAsnGluThrValGluValIleDeSerGluMet

<sup>170</sup> <sup>190</sup> <sup>210</sup> <sup>230</sup>  
 TTTGACCTCAGGAGCCGACCTGCCTACAGACCCGCCCTGAGCTGTACAGCAGGCGCTGCGGGCAGCCTCACCAGCTC  
 PheAspLeuGlnGluProThrCysLeuGlnThrArgLeuGluLeuTyrLysGlnGlyLeuArgGlySerLeuThrLysLeu

<sup>250</sup> <sup>270</sup> <sup>290</sup> <sup>310</sup>  
 AAGGCGCCCTTGACCATGATGGCCAGCCACTACAAGCAGCAGCTGCCCTCCAAACCCGGAAACTTCTCTGTGCAACCCAG  
 LysGlyProLeuThrMetMetAlaSerHisTyrLysGlnHisCysProProThrProGluThrSerCysAlaThrGln

<sup>330</sup> <sup>350</sup> <sup>370</sup>  
 ATTAACCTTTGAAAGTTTCAAGAGAACCTGAAGGACTTTCTGCTGTGTCATCCCCCTTGGACTGCTGGGAGCCAGTC  
 IleIleThrPheGluSerPheLysGluAsnLeuLysAspPheLeuLeuValIleProPheAspCysTrpGluProVal

<sup>390</sup> <sup>410</sup> <sup>430</sup> <sup>450</sup>  
 CAGGAGTGAGACCGCCAGATGAGGCTGGCCAGCCGGGAGCTGCTCTCTCATGAACAAAGAGCTAGAAATCA  
 GlnGluEnd

<sup>470</sup> <sup>490</sup> <sup>510</sup> <sup>530</sup>  
 GGATGCTATCTTGGAGGGACCAAGGGGTGGGCCACAGCCATGGTGGGAGTGGCCCTGGACTGCCCTGCCACACTGA  
 CCTGATACAGTCATCGCAGAGAAATGGATATTTATACAGCAATCTGATTTATATATATTTTAAATAATTAAATTAA

<sup>550</sup> <sup>570</sup> <sup>590</sup> <sup>610</sup>  
 TTTAATTTAATGAC Nucleotide Sequence of MMT-Type Gene (Upper Line) Encoding Human  
 TAAATTACTATTATAC Granulocyte-Macrophage Colony Stimulating Factor Corresponding  
 Amino-Acid Sequence is shown in Lowerline. Mature Protein Begins at Asterisk (\*)

-6 -11 10 \*Sfa.NI 30 50 70  
 AGCTTCTTTGGATTAAGAAGCACCAGCCCGCTCGCCAGCCCGAGCAGCCCTGGGAGCATGTGAATGCCATC  
 Ser Leu Asp Lys Arg Ala Pro Ala Arg Ser Pro Ser Thr Gln Pro Trp Gln His Val Asn Ala Ile

**Fig. 1B.**

82 90 110 130 150  
 CAGGAGCCCTTGGCTCTCTGAACCTGAGTGAAGACATGCTGAGATGAATGAACAGTAGAAGTCACTCAGAAATG  
 Gln Gln Ala Leu Arg Leu Leu Asn Leu Ser Arg Asp Thr Ala Arg Glu Met Asn Gln Thr Val Gln Val Ile Ser Gln Met

20 30 40 50 60 70  
 TTTGACCTTCAGGAGCCGACCTGCTTACACACCCGCTGGAGCTGTACAAGCAGGGCCTGGGGCAGCCTACCAAGCTC  
 Phe Asp Leu Gln Gln Pro Thr Cys Leu Gln Thr Arg Leu Gln Leu Tyr Lys Gln Gly Leu Arg Gly Ser Leu Thr Lys Leu

80 90 100 110 120 130  
 AAGGCCCTTGAACCATGATGGCCAGCCACTACAAGCAGCAGTCCCTCCAAACCCGGAAACTTCTGTGCAACCCAG  
 Lys Gly Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro Pro Thr Pro Gln Thr Ser Cys Ala Thr Gln

140 150 160 170 180 190  
 ATTATCACCTTTGAAAGTTTCAAGAGAACTGAAGGACTTCTGCTGTGTCATCCCTTTGACTGCTGGGAGCCAGCTC  
 Ile Ile Thr Phe Gln Ser Phe Lys Gln Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Gln Pro Val

200 210 220 230 240 250  
 CAGGAGTCAGACCGCCAGATGAGGCTGGCCAGCCGGGAGCTGCTCTCTCATGAACAAAGAGCTAGAAACTCA  
 Gln Gln End

260 270 280 290 300 310  
 GGATGGTCACTTTGGAGGACCAAGGGGTGGGCAACAGCCATGGTGGAGTGGCTGGACTGCCCTGGCCACACTGA  
 CTTGATACAGGCATGGCAGAGAATGGGATATTATATCTGACAAATATCTGATTTATTTATTTATTTATTTATTTATTTA

320 330 340 350 360 370  
 TTTAATTATTTAC Nucleotide sequence of site substituted/Mutated Gene M13HUGMLEU 23 (upper line)  
 TAATTAATTTATTTAC Encoding Analog Human Granulocyte-Macrophage Colony Stimulating Factor The Cor-  
 rection of the sequence is shown in the lower line. Mature Analog Protein Begins at Asterisk (\*)

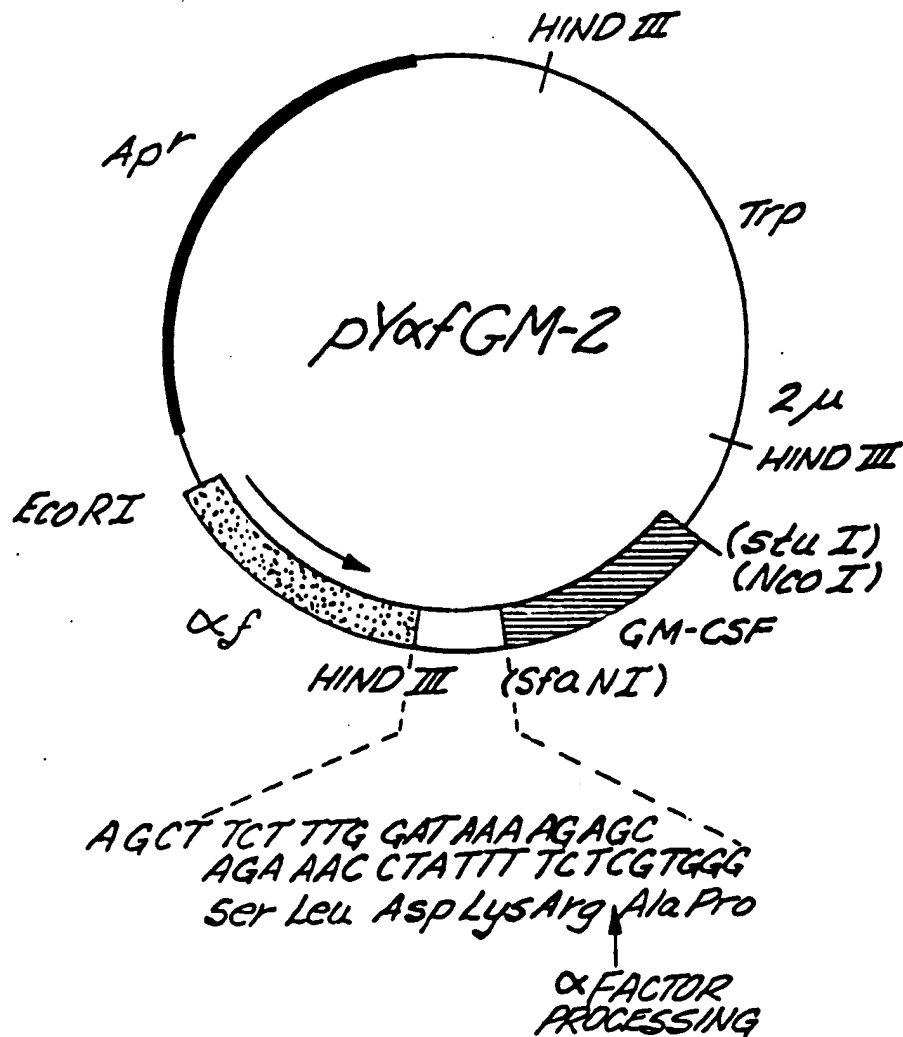
-6 -11 10 \* -SRA.NI 30 50 70  
 AGC TTCTTTGGATTAAGAAGCAGCCGCCGCTCCAGAGCCCGAGACAGCCCTGGAGCATGTGAATGCCATC  
 Ser Leu Asp Lys Arg Ala Pro Ala Arg Ser Pro Ser Pro Ser Thr Gln Pro Trp Gln His Val Asn Ala Ile

# Fig. 16.

CAGGAGGCCGCTCTCTGAGATGAATGAACAGTAGAAGTCATCTCAGAAATG 130  
 Gln Gln Ala Arg Leu Leu Asn Leu Ser Arg Asp Thr Ala Ala Gln Met Asn Gln Thr Val Gln Val Ile Ser Gln Met  
 20 30 40 50 60 70  
 150 170 190 210 230 250 270 290  
 TTGACCTCAGAGCGACCTGCTTACAGACCCCTGGAGCTGTACAAGCAGGGCCTGGCGGGCAGCCCTCACCAGCTC  
 Phe Asp Leu Gln Gln Pro Thr Cys Leu Gln Thr Arg Leu Gln Leu Tyr Lys Gln Gln Gly Ser Leu Thr Lys Leu

230 250 270 290 310 330 350 370 390 410 430 450  
 AAGGGCCCTTGACCATGATGGCCAGCCACTACAAGCAGCCTGCTCCAACCCCGAAACTTCTCTGTGCAACCCAG  
 Lys Gln Pro Leu Thr Met Met Ala Ser His Tyr Lys Gln His Cys Pro Pro Thr Pro Gln Thr Ser Cys Ala Thr Gln  
 310 330 350 370 390 410 430 450  
 ATATCACCTTGAAGTTTCAAGAGAACCTGAAGGACTTCTCTGTCTGTAICCCCTTTGACTGCTGGAGCCAGTC  
 Ile Ile Thr Phe Gln Ser Phe Lys Gln Asn Leu Lys Asp Phe Leu Leu Val Ile Pro Phe Asp Cys Trp Gln Pro Val  
 100 390 410 430 450  
 CAGGAGTCAGACCGCCAGATGAGGCTGCGCCAGCCGGGAGCTGCTCTCTCATGAACAAAGAGCTAGAAACTCA  
 Gln Gln Leu End 470 490 510 530

GGATGCTATCTTGAGGGACCAAGGGGTGGGACAGCCATGGTGGAGTGGCCCTGACTGCTCTGGCCACACTGA 530  
 CCTGATACAGGCATGGCAGAGAAATGGATATTATCTACAAATCTGATTTATATATTAATTTTAAATATTTTATTTTAA 610  
 TTTAATTTAATTGAC Nucleotide Sequence of Site Deleted Mutated Gene M13HUGMΔArg 23 (upper line)  
 Encoding Analog Human Granulocyte-Macrophage Colony Stimulating Factor The Cor-  
 responding Amino Acid Sequence is Shown in the Lower Line. Nucleotide Positions of Deleted

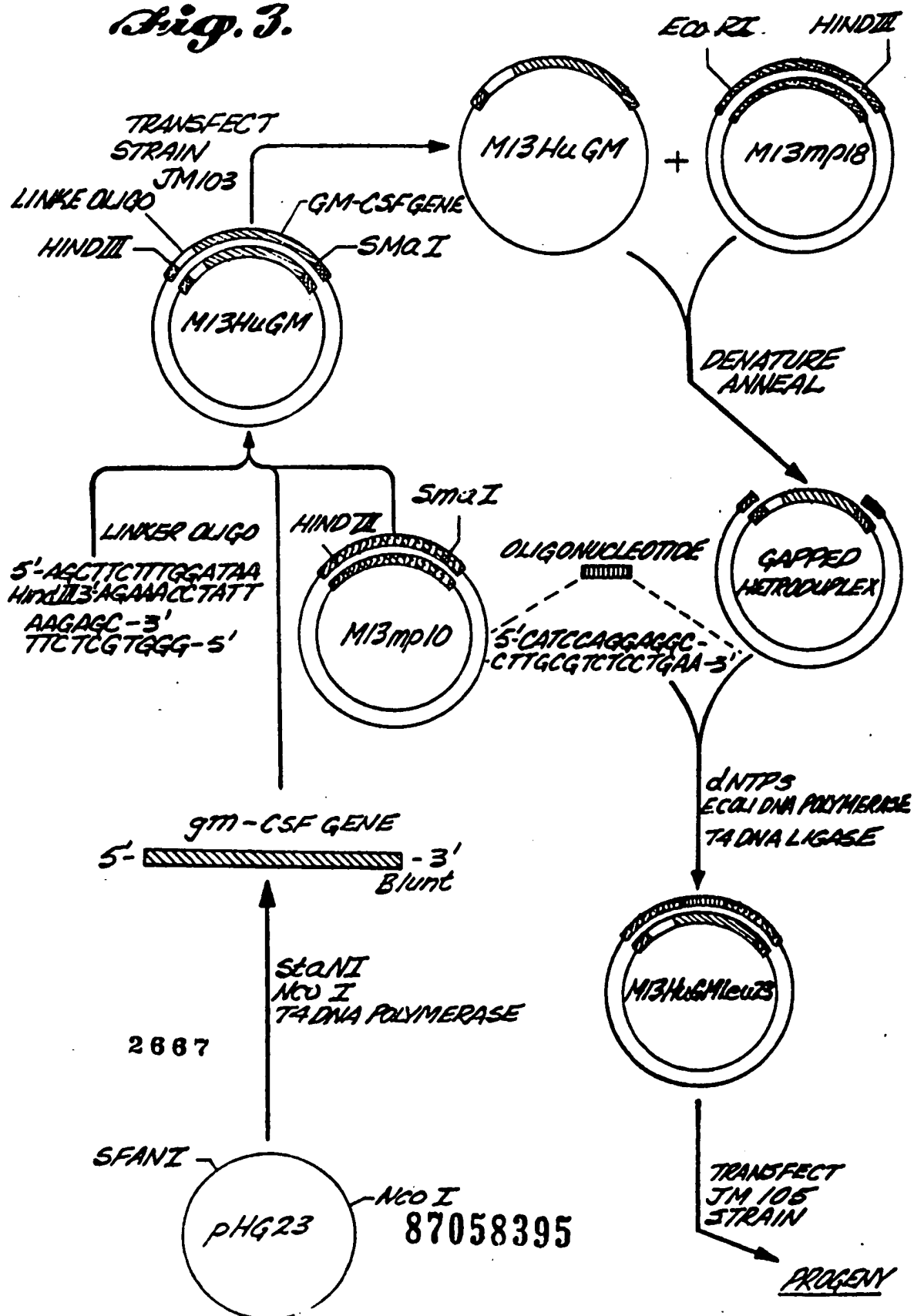


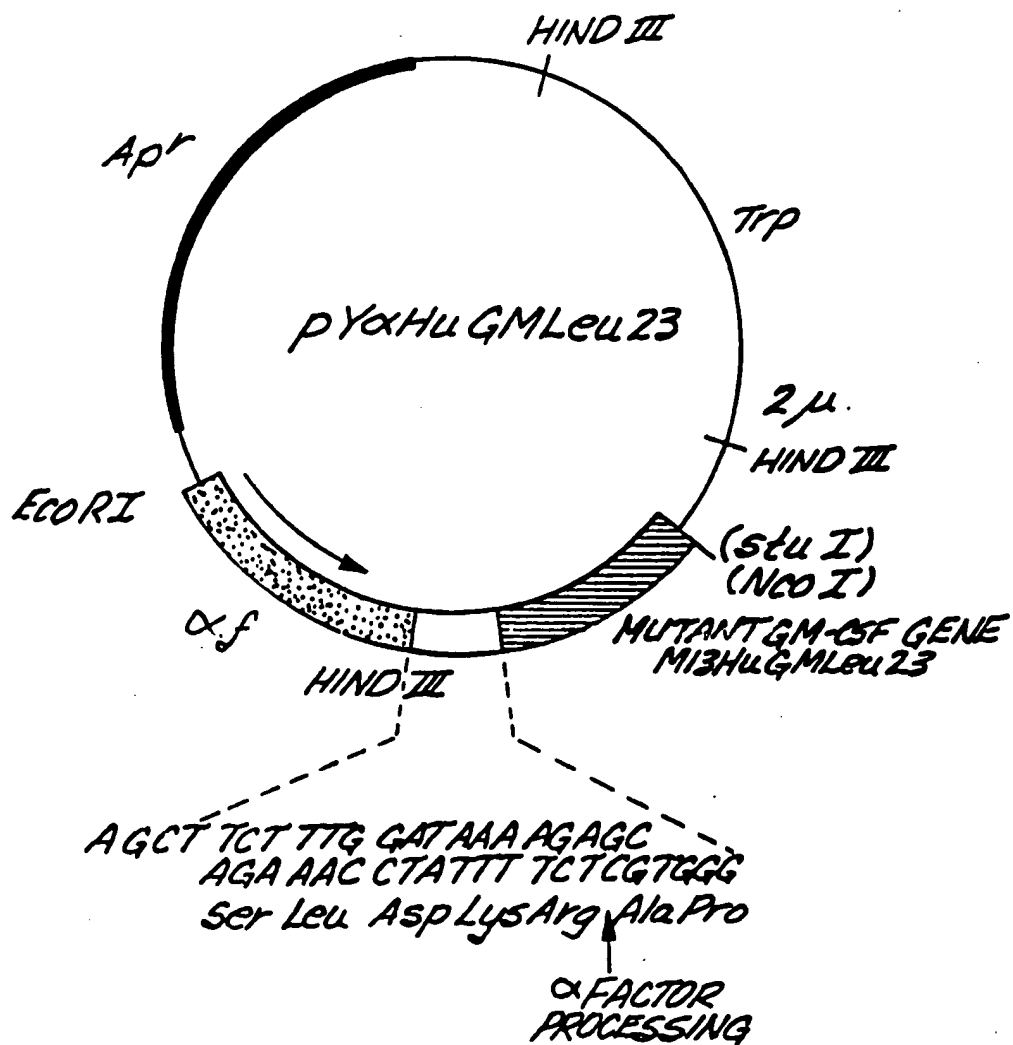
2666

*Fig. 2.*

87058395

Fig. 3.





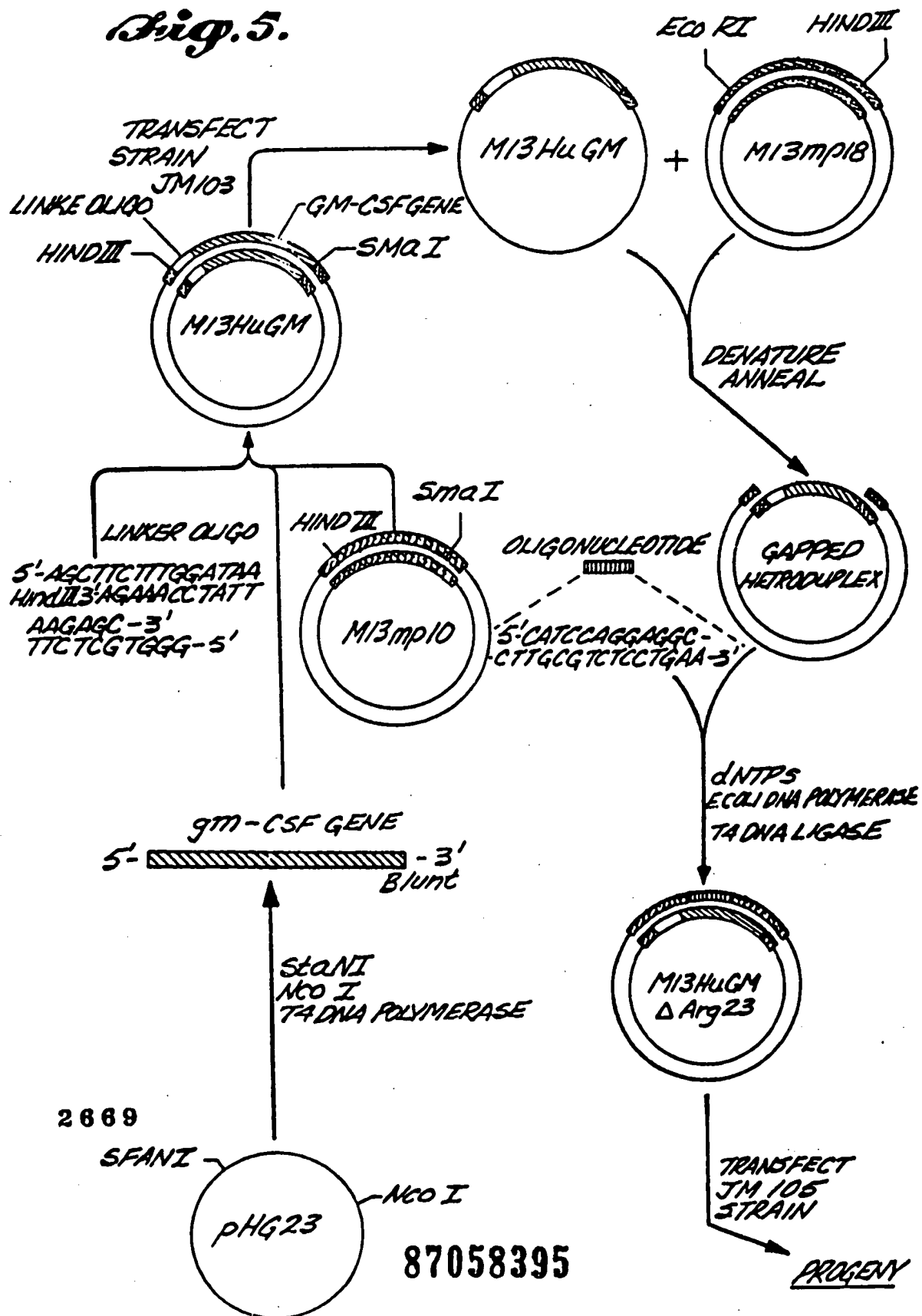
2668

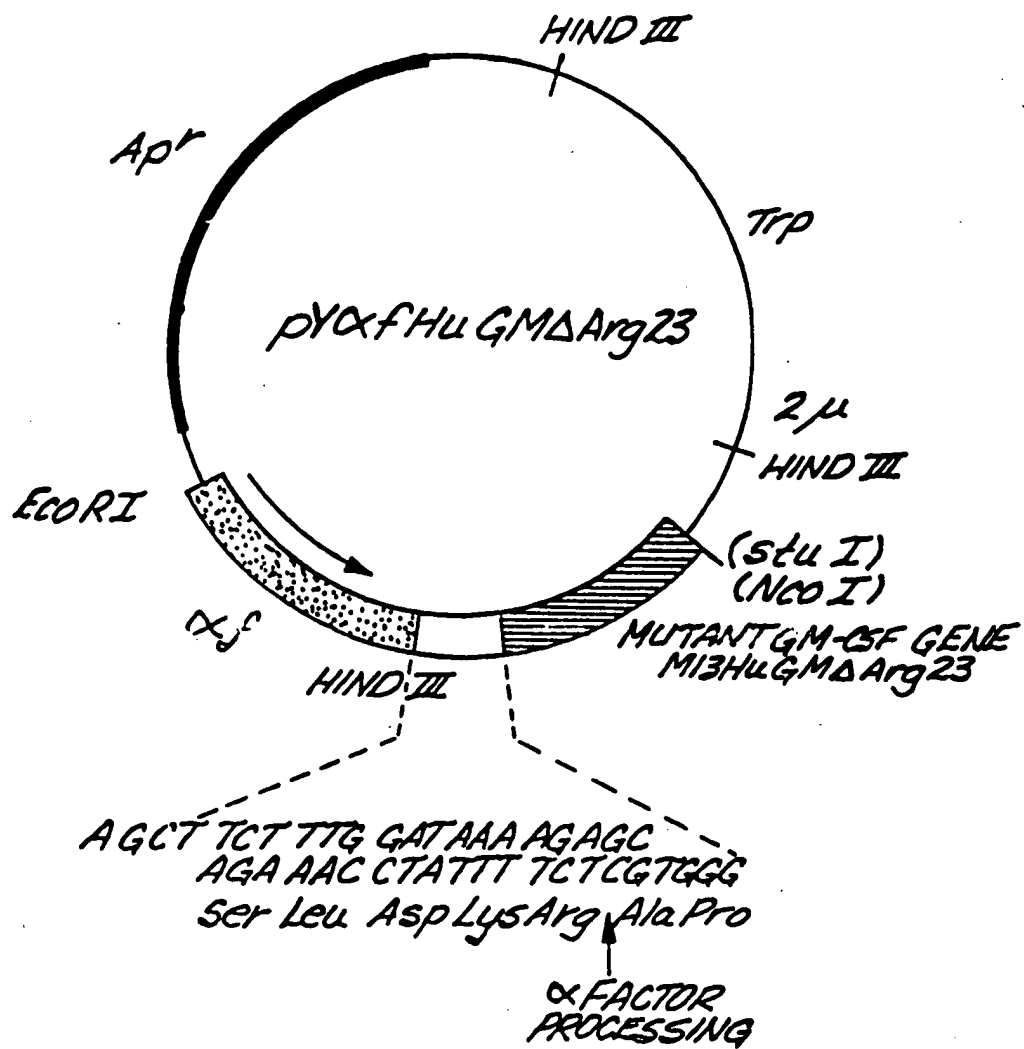
*Fig. 4.*

87058395



Fig. 5.





2670

*Fig. 6.*

87058395